Driving Growth in Insurance With a Big Data Architecture

The SAS® and Cloudera Advantage

Version: 104

Table of Contents

DRIVING GROWTH
IN INSURANCE

WHITE PAPER

Overview

Data has always played a critical role in insurance. The insurance business is built on analyzing data to understand and evaluate risk. The big data era is dramatically changing the insurance industry, prompted by the innovation of new technologies in data consolidation and unified analytics that give insurers new strategic and operational insights into their business.

A recent survey[1] found that 89 percent of senior executives believe big data is changing the way they do business. It enables insurers to select more profitable business strategies, implement more precise pricing, manage the risk portfolio holistically, and improve fraud detection.

Current Data Challenges for Insurers

Fragmented systems, legacy applications, and data silos have long been an impediment for insurers. They are costly to maintain and integrate. Moreover, these systems are not designed to handle the sheer volume, complexity, and types of data that are needed by businesses today. This can include – historical data, web logs, clickstreams, social media, geospatial, traffic, weather and sensor data, data stemming from omni-channel touch points, digital, and connected devices.

As a result, IT departments are experiencing decreased data processing performance, ballooning storage costs, or both. Traditional data management strategies do not scale to effectively govern large data for high-performance analytics. The most common obstacle for companies is that they have too much data and too few IT resources.

Insights from analytics are only as good as the data that is fed into the analytical models and algorithms. Since data and analytics form the backbone in the management of an insurance company, the quality, consistency, and breadth of the data become even more important.

Unlocking the Power of Big Data Requires a Modern Data Architecture.

Leading insurers of all types and sizes are harnessing the power of big data by deploying an Enterprise Data Hub (EDH) powered by Hadoop and combining that with high-performance analytics. An EDH delivers a new paradigm for overcoming data silos. It enables insurers to aggregate and analyze massive amounts of data of all types, in any volume, from any source, for all types of workloads – all from a single platform. Unlike the "rip-and-replace" options of the past, an EDH allows insurers to take advantage of their existing IT investments – from storage technologies and enterprise applications to back-end legacy systems.

Moreover, Hadoop has caught the attention of many organizations searching for better ways to store, process, and analyze large volumes and varieties of data from multiple internal and external sources. Based on a survey by TDWI[3], 88 percent of organizations consider Hadoop a strategic investment:

- It's cost-effective. Hadoop uses lower-cost commodity hardware to reliably store large quantities of data.
- It's scalable. Hadoop provides flexibility to expand out by simply adding more nodes.
- It's adaptable. Dump any type of data in its original formation into Hadoop and apply structure as needed for consuming application.

The combination of an Enterprise Data Hub (EDH) with high-performance analytics is fast becoming a competitive advantage for insurers. With high-performance analytics, insurance companies can prepare, explore, and model multiple scenarios using data volumes never before possible. They can also process complex analytical algorithms faster – quickly delivering better answers to decision makers and key stakeholders across the organization.

A 2014 survey by advisory firm[2], Strategy Meets Action, found that 39 percent of Tier 1 insurers (premiums above $1bn) are investing in big data technologies. That's a significant jump from just 10 percent in 2013.

[1] Accenture: Big Success with Big Data accenture.com/us-en/insight-big-success-big-data.aspx.
[2] Strategy Meets Action: Big Data in Insurance
[3] TDWI Best Practices Report: Integrating Hadoop into Business Intelligence and Data Warehousing, Second Quarter, 2013

Cloudera Enterprise Benefits
Stores and Analyzes Any Type of Data
- Leverage the full power of your data to achieve pervasive analytics, increase business visibility, and reduce costs
- Bring diverse users and application workloads to a single, unified pool of data on common infrastructure; no data movement required

Enterprise Approach
- Compliance-ready perimeter security, authentication, granular authorization, and data protection through encryption and key management
- Enterprise-grade data auditing, data lineage, and data discovery

Industry-Leading Management and Support
- Best-in-class holistic interface that provides end-to-end system management and zero-downtime rolling upgrades
- Open platform ensures easy integration with existing systems
- Open source to achieve stability, continuous innovation, and portability

Solving Complex Big Data Problems

The very foundation of an insurer's business is based on accurately assessing and mitigating risks of all types. This requires the availability of the right data at the right time from multiple sources simultaneously. Moreover, empowered consumers, competitive demands, and escalating external threats continue to dictate that new insights be generated much more rapidly – in many cases in real-time.

An Enterprise Data Hub (EDH) and high-performance analytics for big data are designed to solve complex problems and enhance key areas of operations that directly impact loss-ratios and profitability including:

Claims Fraud
Fraudulent activities are on the rise, and if that behavior is not discovered quickly, it may never be detected. An Enterprise Data Hub (EDH) and high-performance analytics enable insurers to detect and prevent fraud by accessing, blending, and analyzing a variety of data from multiple internal and external sources. Insurers can tap into decades worth of claims and policy data as well as data stemming from social media, weblogs, and adjusters' notes, among others.

Claims Subrogation
It is estimated that on average, 5 percent of claims that should go to subrogation don't. By analyzing unstructured data such as adjusters' claim notes and using text mining techniques, insurers can minimize the number of missed recovery cases by recognizing known and unknown subrogation indicators in the claims information.

Ratemaking and price optimization
With an Enterprise Data Hub (EDH) and high-performance analytics, actuaries can develop and run multiple models quickly to produce model uplift and more accurate pricing, without having to do sampling and testing. In addition, insurers can enable real-time pricing or price optimization based on risk attributes, key demographics, competitors, and customer demand elasticity models.

Catastrophe (CAT) Modeling
Catastrophe losses can have a significant impact on the financial stability of an insurance company. Insurers must evaluate their loss exposure and financial position to meet liquidity requirements, often in a real-time environment. An Enterprise Data Hub (EDH) and high-performance analytics enable insurers to blend and correlate massive amounts of weather, climate, geographic, and risk portfolio data to create more accurate models. By using data visualization, carriers can create geographical risk exposure reports by augmenting existing policy data with geospatial data to assess and monitor loss exposure by geographic region.

Benefits of SAS

High-Performance Analytics

- Get the timely insights needed to make decisions in an ever-shrinking window of opportunity

- Discover precise answers for complex problems

- Seize opportunities for growth that otherwise would remain unrecognized, and achieve better performance

Comprehensive Data Visualization

- Explore all relevant data quickly and easily

- Look at more options, uncover hidden opportunities, identify key relationships and make more precise decisions to drive success faster than ever before

- Self-service, ad hoc visual data discovery and exploration

Seamless Connectivity

- Integrate data stored in Hadoop (Hive tables) with data from other sources

- Access data directly and easily with a native interface

- Support both technical and business users

Customer Insights (360 degree view)

As customer interactions in insurance moves from in-person to digital channels, insurers not only have to react faster, but they must also be able to predict future behavior. With high-performance analytics for big data, insurers can detect changes in customer behavior in real-time during digital interactions. In turn, they can improve customer experiences and make relevant, real-time offers with higher acceptance probabilities.

Mid-size P&C Insurer Realizes Tangible Value

A privately owned Tier 1 UK insurer, with annual revenue in excess of £1.5bn, was experiencing considerable growth in its customer base. It created challenges in its existing data infrastructure. In particular, the company was struggling to utilize the breadth of new data sources available, like telematics, and process its growing data volumes in a timely manner. The insurer was also trying to stem the effects of increasing application and claims fraud. Unstructured data from third-party sources like Experian that could help identify fraud, were not being fully utilized because this insurer struggled to incorporate the data into its models in a timely fashion.

Recognizing that its existing infrastructure wasn't adequate to make use of the growing and diverse data volumes, the insurance company adopted a Hadoop environment. With tight integration between Cloudera's EDH and SAS® Analytics this insurer's big data project delivered significant and tangible results within six months.

A major benefit from this project is that the insurer is now able to provide real-time price optimization. Before the change it took eight weeks between the time the premium rates were calculated and made available to its distribution channels. With the new Cloudera and SAS big data environment this insurer can now analyze tens of millions of quotes in less than a minute and provide real-time pricing updates.

Other benefits that the Cloudera and SAS big data environment delivered include:

- Approximately £5 million reduction in claim costs through better fraud detection and prevention at point-of-quote

- 120 percent increase in policy count over an 18-month period

- 50 percent reduction in customer cancellation rates and increased customer retention at renewal

The SAS and Cloudera Advantage

SAS and Cloudera enable insurers to unleash value from all their data through a proven combination of enterprise-ready storage, processing, data management, and high-performance analytics.
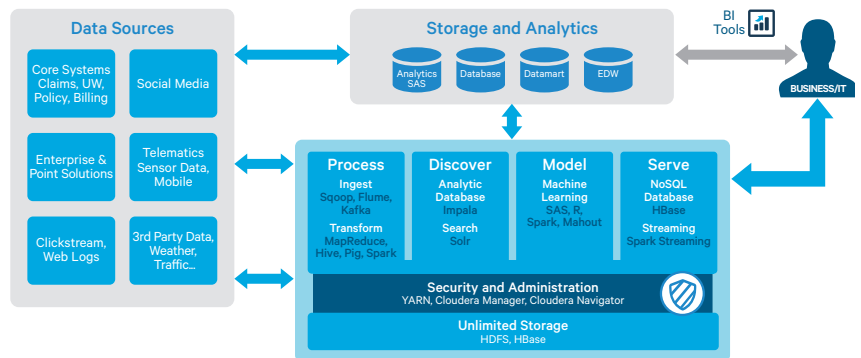
SAS and Cloudera provide a comprehensive infrastructure, solutions and services that can help you:

Improve data access and data preparation for analytics

Cloudera Enterprise powered by Hadoop is a single platform that can serve as a landing place for consolidating and analyzing all your data, in any format, from any source for all types of workloads – all in a highly secure environment. Only Cloudera provides comprehensive, compliance-ready data security and governance throughout the Hadoop cluster, a critical component for insurers working with sensitive, often personal information about customers.

Reduce data movement and replication

This means less time is spent on data preparation and more time on using analytics to drive faster and better decisions. Simplified data preparation saves IT resources and moves the data quality functions into the hands of those who understand the data better. Intuitive web-based user interfaces reduce Hadoop skills gaps and shorten the learning curve. Users can easily create custom data directives (e.g. filters, aggregations and data movement commands) and reuse them when needed. Data can be batch loaded or streamed real-time through Spark.

Deliver faster, analytics-based decisions

The speed of SAS® Analytics is accelerated when running on Cloudera Enterprise by spreading the processing across the nodes in the cluster. In-memory analytics allow you to load data into memory once, then focus on segmenting your data and building multiple models to target groups simultaneously. Guided selection of analytic techniques speed exploration and insights while reducing latencies often associated with model development. Automatically deploy analytical models into Hadoop and score new data using SAS® directly within each node of a Hadoop cluster for fast and efficient processing without data movement.

Easily combine data exploration with predictive modeling

Insurers can discover and evaluate new opportunities for every possible angle. Cloudera Enterprise includes a variety of query frameworks - including Impala, the fastest interactive SQL database on Hadoop – for rapid data discovery. Combined with SAS®, In-memory Analytics, it allows more scenarios and iterations to be run to produce more precise insights. Modelers can quickly test new ideas, try different modeling techniques and refine models on the fly to produce the best results – using data volumes never before possible. Proven, built-for-purpose statistical algorithms and machine-learning techniques deliver precise answers quickly and confidently.

cloudera | **§sas**

About the Authors

Stuart Rose is Global Insurance Marketing Director at SAS. He is responsible for thought leadership and marketing content for applying analytics within the insurance industry. Stuart began his career as an actuary and now has more than 25 years of experience in the insurance industry.

He is a regular contributor to insurance publications and the Analytic Insurer blog. Stuart frequently speaks at insurance conferences and is co-author of the book Executive's Guide to Solvency II. He holds a BSc in mathematical studies from Sheffield University.

Maneeza Malik is responsible for Global Vertical Marketing - Financial Services and Insurance at Cloudera. She has worked with banks and insurers across North and South America, EMEA and APAC for over 15 years. Prior to Cloudera, she spent several years at IBM and ILOG (acquired by IBM). She holds a BCom in marketing and business statistics from Concordia University.

About SAS

SAS is the leader in business analytics software and services, and the largest independent vendor in the business intelligence market. Through innovative solutions, SAS helps customers at more than 75,000 sites improve performance and deliver value by making better decisions faster. Since 1976, SAS has been giving customers around the world THE POWER TO KNOW®. www.sas.com

About Cloudera

Cloudera is revolutionizing enterprise data management by offering the first unified Platform for big data, an enterprise data hub built on Apache Hadoop. Cloudera offers enterprises one place to store, process and analyze all their data, empowering them to extend the value of existing investments while enabling fundamental new ways to derive value from their data. Only Cloudera offers everything needed on a journey to an enterprise data hub, including software for business critical data challenges such as storage, access, management, analysis, security and search. As the leading educator of Hadoop professionals, Cloudera has trained over 40,000 individuals worldwide. Over 1700 partners and a seasoned professional services team help deliver greater time to value. Finally, only Cloudera provides proactive and predictive support to run an enterprise data hub with confidence. Leading organizations in every industry plus top public sector organizations globally run Cloudera in production. www.cloudera.com

cloudera.com

1-888-789-1488 or 1-650-362-0488
Cloudera, Inc. 1001 Page Mill Road, Palo Alto, CA 94304, USA